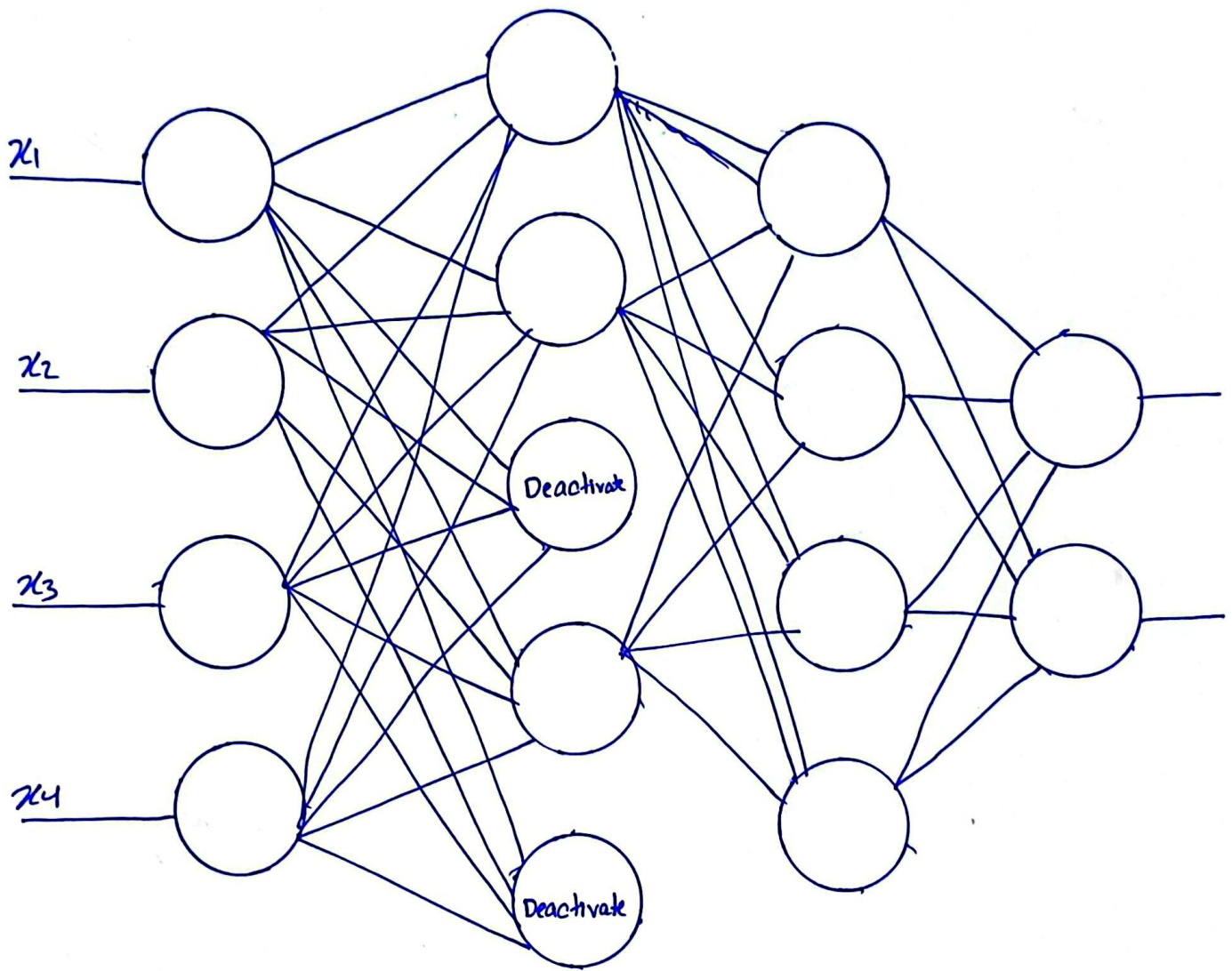


Dropout Layer



Overfitting

Training Accuracy $\Rightarrow 90\%$

Test Accuracy $\Rightarrow 60\%$

Dropout

$$0 \leq p \leq 1$$

$p = 0.5$, 50% of the nodes will be deactivated

Test data

Multiply weights with the initialized probability